Offic de la propriété int ll ctuell du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Prop rty Offi
An Agency of Industry Canada

Bureau canadien des brevets

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

Canadian Patent Office

Certification

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,309,371**, on June 16, 2000, by **THE UNIVERSITY OF BRITISH COLUMBIA**, assignee of Christopher J. Ong, for "Gene Sequence Tag Method".

Tracy Faulkner
Agent certificateur/Certifying Officer

April 22, 2003
Date

Canada

OPIC CIPO

GENE SEQUENCE TAG METHOD

FIELD OF THE INVENTION:

This application relates to a method for indexing a library of genetically altered cells and methods of screening and isolating a particular clone of interest from the library using the high throughput DNA microarray technology. The library is used as a source for identifying and obtaining specifically mutated cells, cell lines derived from the individuallly mutated cells, and cells for use in the production of transgenic non-human animals.

Gene Sequence Tag methodology also relates to an efficient and rapid method for the identification of novel genes, rapid determination of its chromosomal map position and placement of genes on the physical map for the generation of gene transcript maps for eukaryotic genomes and simultaneous generation of gene knock out organisms for in vivo gene function analyses of corresponding genes.

BACKGROUND:

The human genome is estimated to contain approximately 80,000 to 100,000 genes. The age of genomics has resulted in an explosion in the pace of identifying new gene sequence information. However, only a small fraction of these genes have a known role. Nucleotide sequence information alone is insufficient to predict gene function. As the number of gene sequences increases, new global-scale approaches will be required in defining gene function and associating genetic information with human disease.

One of the most powerful methods of revealing gene function, as demonstrated in bacteria, worms, yeast, zebrafish and flies, is the use of forward genetics approaches (from phenotype to genes). Forward genetics entails genome-wide mutagenesis and subsequent screening or selection for mutants affecting defined pathways using phenotypic or molecular assays. State-of-the-art functional genomics employing large scale mutation analyses in lower organisms have proven to be one of the most effective strategies in studying gene function, defining complex genetic pathways and understanding key biological processes. The forward genetics approach is extremely powerful since no assumptions are required about the underlying genes involved.

The mouse is by far the best animal model system for studying human gene function. The mouse model possesses significant advantages for studying human gene function because of its evolutionary relatedness to humans, the ability to rapidly identify homologous genes through regions of genomic synteni, and its similarities with respect to the development of complex tissues and organs. However, functional studies of the mouse has largely been restricted to reverse genetics approaches (from genes to phenotype) which requires a priori assumptions to be made concerning the function and likely phenotype of the mutated gene. The mouse has historically not been amenable to forward genetics approaches because large-scale mutagenesis of the mouse has been cumbersome and expensive due to their large housing requirements, long generation times, high cost of maintenance etc.

Technological advances in *in vitro* culture and maintenance of murine embryonic stem (ES) cells, mouse embryology and molecular biology have provided novel strategies for large scale mutagenesis of the mouse genome. ES cells are derived from the inner cell mass of a 3.5 day p.c. embryo or blastocyst and can be maintained in an undifferentiated pluripotent state under certain culture conditions. ES cells can be genetically manipulated during *in vitro* culture and can subsequently be reintroduced into the mouse embryo by blastocyst microinjection or embryo aggregation techniques. Upon reintroduction into the mouse embryo, ES cells have the potential of contributing to the formation of all tissues of the resulting chimeric mouse. ES cell contribution to germ cells of the reproductive organs results in germline transmission of genetic alterations introduced into the mouse ES cell genome. The mutagenesis of ES cells through gene targeting via homologous recombination and through random means provides a powerful approach for the systematic gene function analyses of the mouse.

The present challenges in mouse functional genomics is to develop systematic, genome-wide mutagenesis and high throughput screening approaches for gene function studies. Large-scale *in vitro* random mutagenesis of ES cell has therefore become a very attractive approach towards mouse functional genomic studies. In order to develop an integrated strategy for large scale mutagenesis of ES cells, two important parameters need to be considered. The first involves the mode of mutagenesis and choice of mutagen. There are currently two widely accepted approaches for mutagenesis. The first involves mutagenesis

through chemical means (such as ENU etc.) or physical means (such as UV or gamma irradiation) and the second involves the use of insertional mutagenesis such as transposon tagging, retroviral integration or gene trap mutagenesis. The insertional mutagenesis approach has many advantages over classical physical or chemical mutagenesis. One of the key advantages of insertional mutagenesis is that it provides a molecular tag for rapid identification and cloning of mutagenized genes whereas chemical and physical mutagenesis approaches require labour intensive physical mapping and genotype scanning techniques to pin point the mutated genes.

The second parameter that needs to be considered is the assay systems for screening or selection of mutants for analyses. Several pre-screening strategies for selecting mutant ES clones bearing mutations affecting interesting genes have been proposed. The strategies can be divided into two broad subsets: phenotype- or expression-based screening approaches and gene sequence-based screening strategies. Phenotype or expression-based screening assays emphasizes the recovery of novel phenotypes. The phenotype-driven route makes no assumptions about the underlying genes involved. This is an effective approach in lower organism to uncovering novel pathways and genes. Examples of current strategies for expression-based screening of ES gene trap clones include screening for LacZ reporter activity in *in vitro* ES-derived cell lineages (Baker et al., *Dev. Biol.* 185(2):201-14 (1997). Kuwano, *R. Zool. Sci.* 13(2):277-83 (1996)) or within ES cell-chimeric whole embryos (Wurst et al., *Genetics*. 139(2):889-99 (1995)).

However, these screening assays are still too slow, non-specific and expensive for effective genome wide mutation screening of the mouse. Furthermore, the current expression-based screens lacks versatility and must be developed and optimized for each biological system being studied.

Gene sequence-based prescreening systems have been proposed for identification of interesting cellular genes disrupted as a result of gene trap mutagenesis in ES cells. In principle, DNA- or sequence-based screens should be faster and less expensive than screens based on cellular or organismal phenotypes. Hicks et al. (*Nat. Gen.* 16: 338-344 (1997)) and Zambrowicz et al. (*Nature* 392: 608-611 (1998)) have championed the production of a library of ES gene trap clones indexed by DNA sequence information. The premise of their

approach is that most mammalian genes will soon be characterized as partial cDNA sequences or EST's. T utilize this sequence information for large-scale functional studies, a library of ES cells bearing unique gene trap integration events would be highly valuable. This resource is proposed to facilitate and expedite the reverse genetics approach for production of gene knock out mice for large scale gene function analyses. However, this industrial-scale strategy requires saturation mutagenesis of the mouse genome and vast capital resources for cloning and sequencing of each individual trapped gene.

Summary of the Invention

The present invention relates to a method for indexing of cells containing insertional gene trap mutation events. The invention allows the prescreening of cells harbouring mutations in genes that are differentially regulated between defined cellular states such as disease and normal states through the use and fabrication of customized DNA microarrays using unique DNA identifiers that index each of the gene trap cell clones.

The present invention also relates to a method for the generation of gene transcript maps of eukaryotic genomes. The Gene Sequence Tag (GST) methodology can also be used for rapid gene identification and concomitant chromosomal localization and genomic sequencing. One of the major goals of the Human Genome Project is to develop efficient methods for identifying genes and for placement of genes on physical maps or sequenced DNA. A human gene transcript map would provide a valuable resource for the rapid determination of genes responsible for both acquired and inherited diseases. The present technology relates to an efficient and high throughput methodology for the identification of novel genes, the rapid determination of its chromosomal map position and for the generation of gene transcript maps for eukaryotic genomes. In addition, the present invention relates to a method for the concomitant generation of mutant cells or organisms.

The current state of the art technologies for the identification of expressed genes and the generation of a cDNA transcript map of the human genome relies on the utilization of Expressed Sequence Tag Methodologies or EST's. The present invention overcomes many of the limitations intrinsic to the EST methodology and greatly enhances the speed

with which exon/gene coding sequences can be subsequently assigned to contigs or the physical map. This assignment provides a necessary first step in understanding the relationship of the genes to both acquired and inherited diseases by providing candidate causative genes for inherited and acquired disorders (such as cancer) and by providing a detailed gene transcript map for rapid regional definition of disease genes.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Shows the GST gene trap vector indicating the key components.

Figure 2. Shows a schematic diagram of the interactions between the ES gene trap system and other core functional genomic technologies.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an approach that integrates a variety of genomics technologies into a cohesive, broad-based and directed strategy that facilitates the efficient functional characterization of large numbers of genes. This experimental approach allows the expansion of the scope of biological investigation from studying single genes/proteins to studying all genes/proteins simultaneously.

The present invention encompasses an integrated mouse functional genomics strategy that combines the power of expression-based screens with the high throughput efficiency of gene-based techniques. The strategy invloves a method called the gene sequence tag (GST) methodology which combines large-scale gene trap mutagenesis and tagging of mouse gene transcripts in ES cells with the high throughput and versatile DNA microarray technology for genome-wide expression analysis. The invention involves the use of DNA microarray technology for indexing of signature DNA fragments corresponding to trapped genes in each embryonic stem cell gene trap clones and screening for identification of differentially-regulated trapped genes. The microarray technology is highly versatile and easily adaptable to many biological systems. The microarray technology is a proven technique with desirable high throughput efficiency and capacity for automation. The GST technique facilitates gene identification, sequencing, chromosomal localization, physical

mapping, gene expression profile analyses and gene disruption in mice; therefore, allowing the expeditious progression of experimental investigation from gene discovery and detection of differential gene expression to production of knock out mice for *in vivo* gene function analysis.

A resource of gene trap events in ES cells are generated utilizing the gene sequence tag (GST) methodology. Gene trapping is performed using the following gene trap DNA construct comprising two functional units (Figure 1). The first functional segment consists of a mutagenic, detectable component that comprises an unpaired splice acceptor sequence fused to an internal ribosomal entry sequence (IRES) linked to the β-galactosidase reporter gene followed by a polyadenylation signal sequence (SA-IRES- β gal-pA). The second functional unit encodes a selectable sequence acquisition module consisting of the mouse phosphoglycerate kinase-1 (PGK) promoter that is known to be actively transcribed in ES cells fused to the puromycin N-acetyl transferase gene followed by an unpaired synthetic consensus splice donor sequence (PGKpuroSD).

Transfection of the GST gene trap DNA construct via electroporation into ES cells results in random integration (the majority of which are single copy vector integration events) into the ES cell genome by illegitimate recombination. The selectable PGKpuroSD gene cassette lacks a polyadenylation signal sequence; therefore, puromycin resistance can only be achieved by splicing into downstream exons and polyadenylation signal sequence of the trapped endogenous gene. The GST gene trap vector not only introduces a molecular tag that provides a handle for subsequent cloning and identification, chromosomal localization and placement onto the physical map of the trapped gene, but also simultaneously generates ES cells bearing mutations in the respective genes that facilitates generation of knock out mice.

Each ES cell gene trap clone obtained using the GST method simultaneously provides rapid access to the following key pieces of information: 1) Partial cDNA gene fragments corresponding to the trapped genes can be cloned by rapid amplification of cDNA ends (3'RACE-PCR). 2) The identity of the novel genes trapped can be determined by obtaining partial gene sequence information through high throughput DNA sequencing of RACE PCR products. 3) The chromosomal localization of the trapped genes can be

identified by fluorescence in situ hybridization (FISH) mapping. 4) The genomic DNA sequence flanking the site of integration can be rapidly cloned and sequenced providing sequence information that will allow for rapid placement of genes on DNA contigs or the physical map. 5) The direct histochemical demonstration of the pattern of gene expression (due to the presence of the LacZ reporter gene) in either chimeras or germline animals produced with ES cells can be attained. 6) In vivo gene function information can be obtained from phenotypic, physiologic and biochemical analyses of ES cell-derived knock out mice and cell lines.

A partial or complete set of randomly genetically altered cells is generated. For example, a library of ES cell gene trap clones is generated by random insertional mutagenesis using the above described GST gene trap DNA vector. Each trapped gene is cloned by 3'RACE-PCR. PCR products are then used for the fabrication of DNA microarrays. An object of the present invention is a DNA microarray of PCR products that index the individual genetically altered cells within a library. Quantitative gene expression analyses using DNA microarray hybridization is subsequently performed in order to identify differentially expressed genes in a variety of model systems. Gene chip hybridization probes derived from test and control cell- or tissue-samples are prepared from defined biological systems, for example: neurodegenerative disease; DNA repair; prostate cancer; adhesion signalling; macrophage activation; immune tolerance and activation; Apoptosis; dendritic cell function and liver regeneration. The advantage of using the DNA microarray technology up front prior to DNA sequencing is that DNA sequencing will be restricted to differentially regulated genes. This represents a huge economical saving.

Based on the literature[4], approx. 300-500 differentially regulated genes per 10 000 trapped genes spotted on a microarray chip are estimated to be identified. A significant cross-regulation of certain gene classes in multiple biological systems is anticipated. For example, genes that are up-regulated in apoptosis, neurodegenerative disease and T cell anergy may be conversely down-modulated in cancer progression, liver regeneration, T cell activation etc. One of the defining characteristics of 'genome scale expression profiling' experiments is that the examination of so many diverse genes gives a perspective on all the processes that simultaneously occur within a model system. The comparison of gene

expression profiles between model systems will provide new insights into the role of genes in the context f multiple processes. Therefore, this approach will identify novel genes and gene families that play common and unique functional roles in multiple pathways and systems.

PCR products from corresponding differentially regulated trapped genes identified by microarray hybridization are used as DNA templates for sequencing. ES gene trap are subsequently selected for chromosomal localization by FISH mapping. Flanking genomic DNA sequence are cloned and subsequently sequenced. Bioinformatic analyses of partial gene sequence information and chromosomal localization is then performed. By comparison of gene sequence and chromosomal position with available public databases, information with respect to whether the trapped genes are novel or known; are part of a gene family, contain known functional domains, etc. is then determined. Based on the results of bioinformatics, specific ES cell clones are for generation of knock out mice and determination of in vivo gene expression pattern. Homozygous mutant mice and cell lines are then used for phenotypic, biochemical and physiologic analyses. Subsequent cycles of gene identification may be performed using hybridization probes derived from mutant mice and cell lines for further rounds of microarray hybridization studies.

It is important to stress that the GST approach is an unbiased, random, genome wide mutation scanning of the entire genome to identify and functionally characterize novel genes in the absence of a priori assumptions about gene sequence and function. The core platform technology has broad applicability for studying a wide variety of biological and physiological processes. This novel functional genomics strategy will initially be utilized to systematically dissect several key biological research areas, namely: neurodegenerative disease, cellular stress, prostate cancer, adhesion signalling, macrophage activation, immunologic tolerance induction, apoptosis, dendritic cell function, and liver regeneration.

The expressed sequence tag or EST method is currently the primary method of choice for large scale gene discovery efforts. While the EST method has identified a vast number of genes, EST data are inherently incomplete. The major drawbacks of the EST technology relate to the fact that ESTs are subject to expression bias and redundancy. The set of genes discovered via ESTs are highly dependent on the tissues/cells from which the

cDNA libraries are prepared and the relative gene abundance. Rare gene transcripts are unlikely to be discovered within EST data sets while abundant transcripts are over represented. The GST methodology overcomes many of the limitations of the EST method for gene discovery and provides an additional advantage of direct access to in vivo gene function information via production of ES cell-derived gene knock out mice. Since gene identification and tagging via the GST method involves the random integration of the gene trap vector into the genome, GST gene trapping will likely provide information about coding regions of most genes that is independent of their transcriptional status. Moreover, since the full complement of genes are equally represented in the genome, the GST method will likely identify a relatively non-redundant repertoire of genes.

EXAMPLE

Generation of ES cell gene trap clones: In the following example, gene trap mutagenesis is performed in J1 ES cells. The J1 ES cell line was chosen for the following reasons: 1) J1 cells are derived from a 129 substrain that has been chosen as the source of genomic DNA for the international mouse genome sequencing project. 2) J1 ES cells were originally derived from an inbred homozygous genotype allowing for easy backcrossing to generate knock out in inbred background also allows for gene knock out to be out crossed onto outbred background with a minimal number of matings.

Gene trapping in ES cells is performed using the GST gene trap DNA construct described above comprising a mutagenic, detectable component (SA-IRES-βgal-pA) and a selectable sequence acquisition module (PGKpuroSD). The SA-IRES-βgal-pA component generates a fusion transcript with the endogenous target gene. Previous data have established that most gene trap events containing SA-bgal-pA results in a null allele (Zambrowicz et al.(1998); Skarnes et al., *Genes Dev.* 6(6): 903-918 (1992)). Moreover, the expression of the reporter βgal genes is under the control of the endogenous promoter. The pattern of LacZ activity, therefore, mimics that of the endogenous gene allowing for histological assessment of in vivo gene expression pattern. The internal ribosomal entry sequence (IRES) allows for reporter gene translation independently of the reading frame of the splice junction. The PGKpuroSD component of the GST vector results in expression of

the puromycin resistance gene as fusion transcripts with the 3' end containing downstream exons and the polyadenylation signal of tagged genes. This fusi n transcript allows for the identification of the trapped genes by 3' RACE-PCR in undifferentiated ES cells, even if the genes are not expressed in ES cells.

Host cells are transformed by any of the well-known methods, selected as being suitable for the particular cell type. Electroporation or calcium phosphate mediated transfection are suitable for mammalian cells. A preferred method known for ES cells is electroporation.

A library of GST gene trap ES cell clones each harbouring mutations in unique genes are generated using the GST gene trap DNA vector.

Each ES cell clone has variable cell numbers and growth rates per well after colony isolation. In order to normalize the ES cell numbers per well after clone isolation, ES cells are trypsinized and split into two plates after a few days of culture. One plate is used to determine cell number using an MTT based assay that is detected using an ELISA plate reader. The ELISA Microplate Autoreader EL311 is employed. Using the Bioworks software, discontinuous samples are split by merely supplying a file containing cell number data in a comma-delimited format which is easily be exported from Excel. ES cell clones at varying concentrations in the source plate are individually replated by the automated Biomek 2000 resulting in consolidation of clones having similar concentrations in the destination 96 well plates. After a couple days of culture, three replica 96-well plates are generated using the Biomek 2000 workstation. Two replica plates of ES cells are cryo-preserved using an improved 96-well plate freezing protocol for ES cells that allows long-term storage (Udy and Evans, *Biotechniques*. 17(5):887-94 (1994); Ure et al. *Trends in Genetics*. 8(1):6 (1992); Chan and Evans, *Trends in Genetics*. 7(3):76 (1991)). All plates are barcoded with unique identifiers.

A third replica plate of cells are used for isolation of total polyA mRNA as templates for reverse transcription (RT) and 3'RACE-PCR (3' rapid amplification of cDNA ends-polymerase chain reaction) RNA from ES clones is isolated using a rapid, automated magnetic bead-based mRNA isolation procedure. The Dynal mRNA Direct protocol is automated using a Beckman Biomek 2000 robotic workstation that is adapted with a low-

cost magnetic plate (Dynal XS-96T) located on the work surface of the Biomek 2000. The 96-well plates containing ES cell clones are processed automatically on the Bi mek 2000. Automated RNA extraction is linked with thermocycling by integration of a PTC-200-MJResearch thermocycler (with a robotic lid) adjacent to the Biomek 2000. The Bioworks software program is capable of automated control setup and activation of the RT and 3' RACE PCR reactions using universal primers in the thermocycler. After the PCR run, PCR products are transferred directly for PCR purification using a magnetic bead based procedure called solid-phase reversible immobilization (SPRI) that is compatible with our proposed robotics system. SPRI is an economical, fast, simple and automatable method for the purification of PCR products (96-well format) developed by the Whitehead Institute for Genomics Research. This method has been the method-of-choice used by the Whitehead Institute for Genomics Research for isolation of PCR products for subsequent DNA sequencing.

Purified 3'RACE PCR DNA fragments are then used in preparation of customized high density complementary DNA microarrays.

Preparation of DNA Microarrays: Microarray technology is used to identify and characterize the patterns of gene expression and to identify differentially expressed tagged genes. DNA microarray technology are generally performed on two main types of solid substrates: glass microarrays containing 10,000 to 40,000 DNA spots and nylon membranes containing 1000 to 15,000 DNA spots. Glass slides have several advantages as described (Southern et al., *Nat. Gen. Suppl.* 21:5-9 (1999)) 1) Target DNA is coupled covalently onto treated glass surface. 2) Glass can withstand high temperature and high ionic wash solutions and is non-porous so hybridization volumes can be kept to a minimum which enhances the kinetics of hybridization. 3) Glass has virtually no auto fluorescence and very low non specific probe binding which allows very low signals to be quantitated. 4) Two or more probes can be labelled with different fluorochromes and hybridized together to detect differential hybridization.

There are two aspects of microarray technology: array spotters (robots) and array scanners. Many different prototypes have been recently described each with their reported

advantages (Bowtell, *Nat. Gen. Suppl.* 21: 25-32 (1999)). This technology is advancing rapidly. DNA array spotters are presently available that have the capacity to spot upto 44,000 spots per standard slide (20 mm X 50 mm). The SDDC-1 DNA arrayer supplied by Engineering Services Inc., (Toronto, Ontario) as this instrument has been designed for the production of DNA arrays on glass slides.

Microarrays are prepared by spotting PCR derived DNA products each representing a single gene tag event as described above. The first stage will involve the spotting of 10,000 DNA targets onto a 20 mm x 20 mm area. Target DNA will be prepared and stored in master microtitre plates as described above. Positive controls to be spotted will be 15 housekeeping genes, plasmid DNA, genomic DNA, and 40 spots of GFP DNA. Single stranded cDNA probes will be synthesized from 5 ug of total RNA using reverse transcriptase in reactions containing oligo d(T) primers, deoxynucleotides and either Cy3-dUTP or Cy5-dUTP. Prior to labeling, the RNA population will be spiked with 1ug of GFP RNA produced by in vitro transcription of a plasmid clone with a T3/T7 RNA polymerase initiation signal. This internal control serves to normalize labeling efficiency between RNA preps, to confirm grid location, and measure uniformity of hybridization across the array. Following reverse transcription, RNA will be degraded by treatment with alkali and heat, and fluorescently labeled cDNA purified using Qiagen DNA purification columns. Equivalent amounts of labeled cDNA probes will be combined and hybridized to the microarray under a glass cover slip at 65 C for 8 hours. Slides will be washed at high stringency, dried and scanned for fluorescence.

The microarrays are scanned for fluorescence using the Molecular Dynamics Avalanche scanner with lasers specific for the fluorescence of these probes. The Avalanche is reported to have attomole sensitivity and was designed specifically for use with microarrays.

DNA Sequencing Component: The tagged genes that display changes in expression in different disease states are sequenced using an Applied Biosystems model 373XL automated laser sequencer is used; under optimal conditions the accuracy of the machine is 99% or higher up to about 350 base pairs.

FISH mapping for determining chromosomal location: The enormous advantage of fluorescence *in situ* hybridization (FISH) derives from its unique ability to directly couple cytogenetic and molecular information. Major advances in labelling methods for FISH[8], imaging technologies[9] and our abilities to manipulate eukaryotic chromatin 6 have resulted in FISH mapping methods with a high degree of reproducibility and efficiency[10]. Mapping DNA fragments of 10kb or larger contained in phages, plasmids, cosmids or YACs, using the current FISH technology has been very effective in identifying the chromosomal sublocalization of many novel genes in both human and murine metaphase cells. FISH involves labelling a cosmid, phage, plasmid or BAC/PAC clone with a nonisotopic tag, such as biotin or digoxigenin 9 and the labelled probe is then hybridized to metaphase spreads and the fluorescent signals are detected at the site of hybridization to homologous sequences at one chromosome band location.

A universal probe consisting of the approx. 10 kb Gene Sequence Tag DNA vector is used for mapping experiments of all GST integration events in ES clones identified. The use of a universal probe allows efficient sample throughput. The probes are labelled with biotin-14-dUTP or digoxigenin (DIG)-14-dUTP by nick translation using 'in-house' labelling kits containing reagents that have been optimized and batch-tested for FISH. Duplicate slides are run for each probe, and on average it takes 2-4 FISH laboratory experiments to obtain signals adequate to complete the mapping using 4',6-Diamidin-2-phenylindol-dihydrochloride (DAPI) banding. Since characteristics of universal probe hybridization to ES cell chromosomal DNA is optimized, the determination of chromosomal localization by FISH is more efficient and high-throughput. The chromosome position is confirmed independently by two cytogeneticists. Images are captured as TIFF files converted to JPEG format and subsequently analysed.

GST database: The data gathered from the DNA sequencing and FISH treatment of the ES cell gene tag clones will be compiled in a GST database.

Gene Knock Out Chimeric Mouse Generation: Gene Knock Out Core chimeric mice will be generated from targeted ES cell clones. ES cells harboring trapped genes will be cultivated in an area that is separate from tissue culture involving other types of cell lines that potentially contain murine pathogens, and these will also be spot-checked for infection by MAP testing, prior to introduction into mouse blastocysts.

Weekly microinjections of one to two different ES cell clones into blastocysts for the production of chimeric mice will be performed (approximately 4 and 6 chimeric mice will be produced per ES cell clone). Chimeric mice will be held within the Core Facility until germline transmission is achieved, then mouse lines will be distributed to host labs having access to local animal facilities for further breeding, generation of homozygous mutations, and for specific phenotypic analysis.

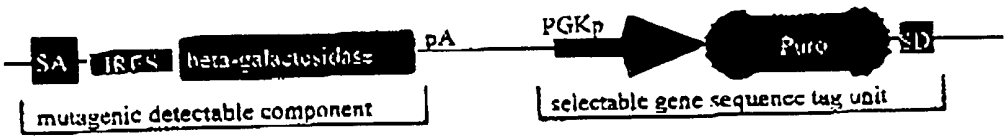

Figure 1. GST gene trap vector. SA – splice acceptor sequence; IRES – internal ribosomal entry sequence; pA – poly adenylation signal sequence; PGKp – phosphoglycerate kinase-1 promoter; Puro – Puromycin resistance gene; SD – splice donor sequence

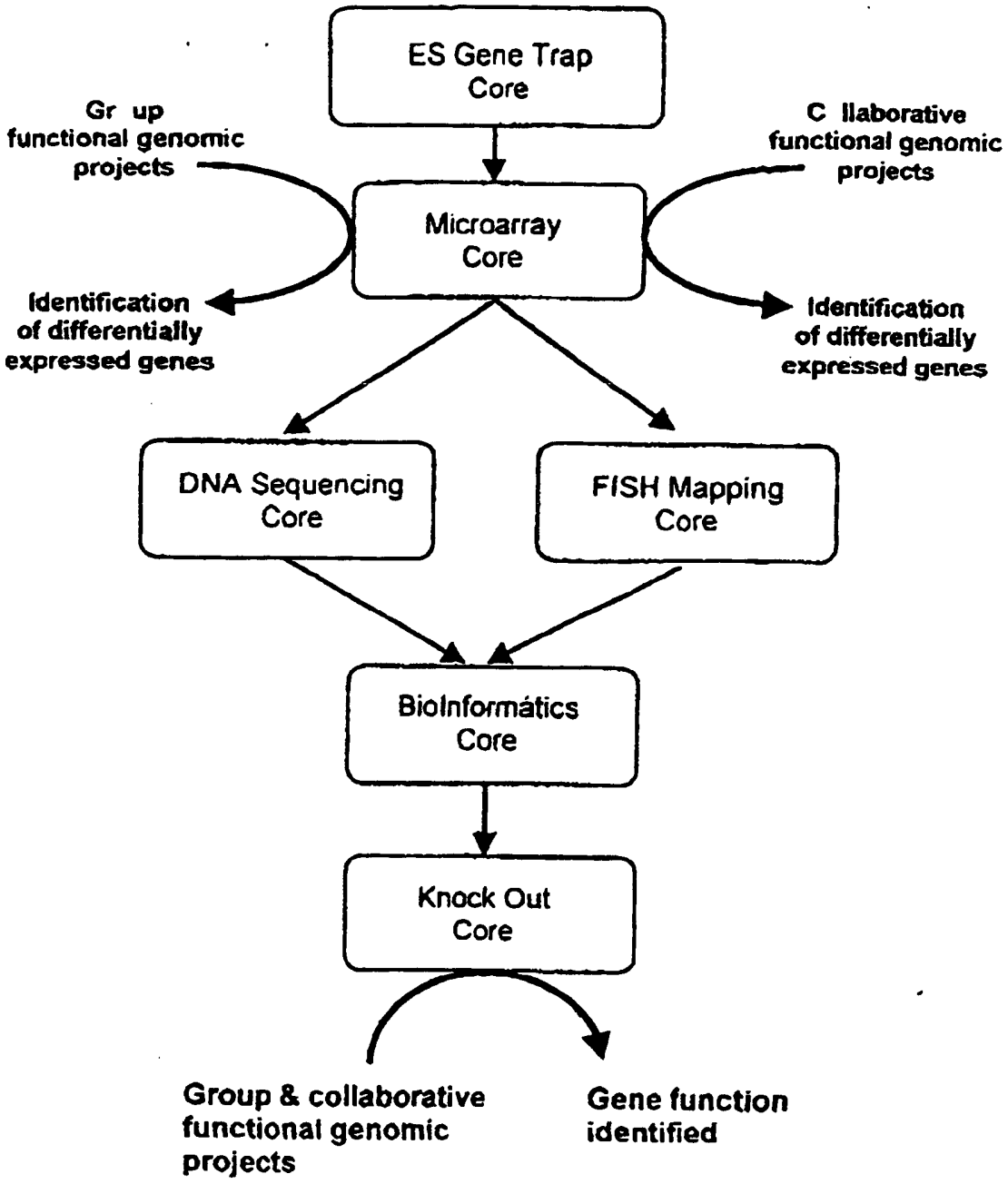

Figure 2 Interaction among ES gene trap core components, group functional genomic projects and collaborations with outside research groups.